

University formative process: quality of teaching versus performance indicators

Stefania Mignani Silvia Cagnone

Department of Statistics, University of Bologna

E-mail: stefania.mignani@unibo.it; silvia.cagnone@unibo.it

Summary: In this work we explore the performance of the Faculties of the University of Bologna by comparing student satisfaction towards the courses attended within each Faculty and performance indicators published by the Evaluation Committee of the University of Bologna. The first source of data is analyzed by a multilevel model. The second source of data is modeled by a factor analysis. The comparison between the results highlight interesting aspects concerning all the Faculties.

Keywords: University evaluation, multilevel models, factor analysis.

1. Introduction

The evaluation of the university formative process and various facets of related quality assessment have recently received a growing attention by Italian Government Institutions. The rising interest of policy makers to systematic procedures for evaluating the university activities, like teaching, research, facilities for students etc., has increased the need of rigorous methodological tools (Minnelli *et al.*, 2005).

In this context the use of proper indicators to measure the efficacy and efficiency of the actions is relevant for a process evaluation based on common aims and largely accepted by Institutions (EU Commission, 2007). In this regard statistical methods can be helpful to define right procedures for an objective and consistent assessment of university system (Bianconcini *et al.*, 2007).

Following the indication of policy makers the Evaluation Committee of Bologna University publishes two reports in each academic year in order to illustrate some results on the activities regarding research and teaching. Some indicators of performances are built to describe the main outcomes.

As for the evaluation of teaching, the data analyzed by the Committee come from two different sources: the Data Warehouse and the Statistical Observatory. The first source contains information on the performance of each student during the academic year: passed exams, mark of each exam, formative credit obtained, and so on. The Statistical Observatory collects and elaborates data coming from student judgment on the quality of teaching (SRT, Student's Rating of Teaching).

These two kinds of data define two different frameworks of evaluation, the first one based on objective measures and the second one based on subjective information related to the perceived quality. A joint overall analysis of all these aspects could achieve reliable results for the definition of benchmark indicators that allows to rank the Faculties.

In this paper we present a comparison of the results of analysis based on the objective indicators of performances and the analysis of student judgments. The comparison is carried out on aggregate data at Faculty level. In Section 2, after a brief description of the data on teaching quality, a multilevel analysis is performed. It allows to analyze the heterogeneity of student judgements for ranking the Faculties on the basis of satisfaction level. Section 3 is focused on the information coming from the Data warehouse. The core of this section is the definition of one or more synthesized indicators of performance of each Faculty for a given academic year. At this aim a statistical method for dimension reduction is applied. Finally Section 4 shows the comparison of the two analyses.

2. Student rating of teaching

Satisfaction analysis aims at evaluating the performance of the Faculties from the student point of view, that is students are asked to judge several aspects about teaching and services related to their university course. The data set object of analysis is described in the following section.

2.1. Descriptive analysis

The data used for the satisfaction analysis refers to students enrolled at the first year in the academic year 2004-2005 and have been collected through questionnaires concerning student satisfaction towards all the courses attended in the 23 Faculties of the University of Bologna (NVA-Università di Bologna, 2007). The questionnaires contain 23 items related to different aspects of teaching and facilities, that can be summarized in 5 different macro areas:

- Organization of the course of studies (items 4-5)
- Organization of the course analyzed (items 6-12)
- Teaching activities (items 13-21, except item 18 that concerns aspects out of the interests of this study)
- Infrastructures (items 22-23)
- Interests and satisfaction (items 24-27)

Permitted responses to the questions are "definitely not", "more not than yes", "more yes than not", "definitely yes" (that is a Likert 4 points scale). These items have not been included in the analysis. The data available are the cumulative percentages of "more yes than not" and "definitely yes" for all the items, for the Overall University and for each Faculty.

A first descriptive analysis has been carried out in order to understand the behaviour of the Faculties and of the items separately.

In Figure 1 the box plots of the percentage of satisfaction for the Faculties are reported. We can observe that satisfaction presents a similar behaviour among Faculties, the median value being around 0.8 and the variability greater than 0.6 and less than 1. However "Preservation of the Cultural Heritage" shows the highest median value whereas "Engineering" has a lower and more variable satisfaction than the others.

Also in this case we can notice that the median values are quite high for almost all the items and the variability is quite low.

Figure 1. Boxplots of the percentage of satisfaction per Faculty

Figure 2. Boxplots of the percentage of satisfaction per item

However items 4 and 5 related to the organization of the courses of study present a slightly different behaviour since the median percentage of satisfaction is moderately low.

In Figure 2 the box plots of the percentage of satisfaction related to each item are reported for the overall data set.

It would be useful to include in the analysis information regarding the course of study attended (for example, as covariate) but, unfortunately, this information is not available in the data set analyzed.

2.2. The multilevel model

In order to analyze the heterogeneity among Faculties, a two stage multilevel model is used. It first takes into account the variability among different items (first level units) within each Faculty and then takes into account the variability between Faculties (second level units). The two-level model can be summarized as follows:

1. First level

the variability among items within each Faculty is modeled by the

$$p_{jk} \sim \text{logit}^{-1}(\alpha_k + \varepsilon_{jk}), \quad \varepsilon_{jk} \sim N(0, \sigma_1^2) \quad (1)$$

where p_{jk} is the probability (percentage) of satisfied for the item j ($j = 1, \dots, J$) in the Faculty k ($k = 1, \dots, K$), α_k is the random effect related to each Faculty and ε_{jk} is the error term of the item j in the Faculty k . The error term allows to model the over-dispersion that can happens when logistic regression is applied to count data (Gelman and Hill, 2007). In more detail we assume that $\sigma_{k1}^2 = \sigma_1^2$ i.e. the variability to be equal among items within Faculties. If $\sigma_1^2 = 0$ the data are not over-dispersed.

2. Second level

the variability between Faculties is modelled through the linear model

$$\alpha_k = \alpha + \delta_k, \quad \delta_k \sim N(0, \sigma_2^2). \quad (2)$$

The random effect associated to each Faculty is modeled around the University average α and variability σ_2^2 .

In the following section we show the results of the estimation of the multilevel model to the overall data set.

2.3. Results

The analysis has been conducted by using the package *lme4* of the software *R* (Gelman and Hill, 2007). The results of the estimated model are reported in Table 1.

Table 1. Parameter estimates and standard errors

Parameter	Estimates	St Deviations
α	1.49	0.06
$\text{logit}^{-1}(\alpha)$	0.82	-
σ_1^2	0.06	0.24
σ_2^2	0.59	0.77

We can observe the estimated average probability (that is $\text{logit}^{-1}(\alpha)$) is equal to 0.82, indicating that the global average satisfaction is quite high for the overall University. Moreover, the variability among Faculties (equal to 0.59) is quite low and not significant, the values of α_k ranging from 1.2 to 2.0. Also the variability among items is very low ($\sigma_1^2 = 0.06$) and not significant, to indicate that there is no over dispersion in the data.

Thus, from these results we can affirm that in general students' satisfaction in the University of Bologna is quite high and this attitude is not very different among Faculties and among items. However, it can be interesting to look at the behavior of each Faculty in terms of satisfaction compared with the global average satisfaction so that we can rank the Faculties from the best (in terms of perceived satisfaction) to the worst. In this regard, in Figure 3 the values of the estimated percentages associated to the Faculties are reported. The solid line represents the estimated percentage of the overall University. The highest estimated percentage is

associated to "Psychology" ($\text{logit}^{-1}(\alpha) = 0.86$) whereas the lowest estimated percentage are associated to "Engineering", "Engineering 2" and "Medicine" ($\text{logit}^{-1}(\alpha) = 0.77$).

Figure 3. Estimated percentage of satisfaction per item

However, these results concern only the perceived quality of teaching. In the next section the student performance analysis is described. This analysis offers interesting suggestions to compare the Faculties.

3. Performance analysis

The data set used for the Performance analysis comes from the Data Warehouse of Bologna University. It contains a lot of information on students' performance.

Table 2. Descriptive statistics for the academic year 2004-2005

	% regular students	% students who take exams	Average of credits	Average of marks
Agriculture	79	72	31	25.4
Architecture	90	87	44	26.7
Arts and Humanities	82	79	33	27.3
Economics FORLÌ	84	80	38	25.6
Economics RIMINI	84	81	34	23.6
Economics BOLOGNA	85	82	33	24.0
Education Sciences	85	81	32	27.6
Engineering	91	88	36	24.9
Engineering 2	83	86	31	24.8
Exercise and Sports Sciences	87	85	32	25.6
Foreign Languages and Literature	82	82	30	27.0
Industrial Chemistry	84	81	46	26.4
Interpreters and Translators	96	97	49	27.0
Law	86	80	35	25.6
Math, Physics and Natural Sciences	83	77	33	25.5
Medicine	94	92	40	25.7
Pharmacy	89	87	29	25.1
Political Sciences	85	79	35	24.6
Political Sciences FORLÌ	85	78	45	25.6
Preservation of Cultural Heritage	76	71	27	27.8
Psychology	96	93	50	26.3
Statistical Sciences	77	74	42	26.0
Veterinary Medicine	92	85	29	25.6

In particular, in order to compare the results of this analysis with the previous one, we considered four aggregate indicators on the performance of students enrolled at the first year (academic year 2004-2005) for each Faculty: i) the percentage of students that paid the second instalment of university fees; ii) the percentage of students who passed examinations and gained course credits; iii) the number of credits that students achieve on average; iv) the average of marks. Table 2 shows the descriptive analysis on the data set considered.

We can notice that "Interprets and Translators" and "Psychology" are the Faculties with the best performances concerning regular course of study (high percentage of students who have paid the second instalment of university fees and high percentage of students who have passed examinations). The Faculties with the worst performances are "Preservation of Cultural Heritage" and "Statistical Science". Considering the average of credits and marks "Interprets and Translators" and "Psychology" achieve the best performances again, but also "Industrial Chemistry" and "Political Sciences" have good results. With respect to the marks average, the highest values are shown by "Preservation of Cultural Heritage", "Education Sciences" and "Foreign Languages and Literature" whereas the lowest values refer to the two "Engineering" Faculties and to "Economics".

In order to define proper indicators that summarize the overall information of these aspects, we applied a standard statistical technique for dimension reduction, the factor analysis, that helps to find latent dimensions (Bartholomew *et al.*, 2002; Skrondal and Rabe-Hesketh, 2004). The amount of variance explained by the first two factors is about 70% of the overall variability. We can therefore consider only two dimensions for representing data: the first factor could be identified as an indicator of regular path of study and the second one as an indicator of achievement in learning.

Figure 4 shows the position of each Faculty with respect to the two indicators: the horizontal axis stands for regular path of studies and the vertical axis stands for the achievement. In detail, positive and high values of the two axes identify a good performance for the two indicators, whereas negative and high values indicate bad performance. The negative values of horizontal axis denote a unsatisfied behaviour in terms of

regular studies, negative values of vertical axis indicate poor results of learning.

Figure 4. Faculties plotted on the first two latent factors

The 23 Faculties present different behaviours as shown by the clear dispersion of the points on the whole plane. Only four Faculties reach good results for both the indicators whereas seven Faculties show bad performances. We notice that "Interprets and Translators" is the Faculty with the best overall performance differently from the Faculty of Economics–Rimini. "Preservation of Cultural Heritage" has good results in terms of the exams but poor results for the regular path of studies. On the opposite side of the graph there is "Engineering" with negative performance on marks but good behaviour on regularity of studies.

4. Comparison between the analysis

These indications suggest the opportunity of making a comparison between the objective indicators and the results of the analysis of student satisfaction. The comparison could highlight some contradictory aspects and could point out some strategies for improving the formative process, the conditions of learning and the facilities.

Table 3. Comparison between MLA and PA

Ranked Faculties	MLA	PA
Foreign Languages and Literature	+	B G
Pharmacy	+	G B
Engineering 2	+	B B
Math, Physics and Natural Science	+	B B
Architecture	+	G G
Political Sciences FORLI	-	B G
Engineering	+	G B
Interprets and Translators	-	G G
Agriculture	+	B B
Medicine	+	G B
Economics FORLI	-	B B
Economics RIMINI	+	B B
Preservation of Cultural Heritage	-	B G
Statistical Sciences	-	B G
Political Sciences	-	B B
Arts and Humanities	-	B G
Veterinary Medicines	-	G B
Law	-	B B
Economics BOLOGNA	+	B B
Exercise and Sports Sciences	+	G B
Education Sciences	-	BG
Industrial Chemistry	-	G G
Psychology	-	G G

In Table 3 Faculty ranking based on the results of both multilevel analysis (MLA) and performance analysis (PA) are shown simultaneously.

In the first column the Faculties are in ascending order with respect to the overall satisfaction. The results of the multilevel model (ML) are shown in the second column where the Faculties with estimated satisfaction above the mean are indicated by a plus sign while the Faculties with values under the mean with a minus sign. As for the results of the analysis of objective indicators, reported in the third column, the Faculties with good performance for both the factors are indicated with GG (see for example, "Psychology"), those with bad performance on regular path of studies and high performance on achievement have signed with BG, those with poor performance on both the factors have BB (see for example "Agriculture"), and finally Faculties with good performance in terms of regular studies but bad results on achievement have GB.

At first glance, results from the various analyses do not fully agree. The Faculty of Psychology has the highest value of overall satisfaction and its estimated value is over the mean. These results are consistent with those obtained in the performance analysis. On the opposite side there is "Engineering 2" with bad results for both satisfaction, observed and estimated, and performance. Also in this case the two kinds of analyses give similar indications.

Moreover, the Faculty of Economics in Bologna received good judgements on satisfaction but the estimated value is under the mean and the results for objective indicators are negative. "Interpret and Translators" does not present a high level of observed satisfaction, but the estimated satisfaction is over the mean and the performance analysis shows a positive behaviour.

Finally, the Faculty of Agriculture has good results for objective analysis but the observed and estimated satisfaction are low.

These preliminary considerations highlight the need of deeper studies on both the aspects. Teaching satisfaction and services not always corresponds to good success in the path of studies. In particular related analyses could be carried out in order to consider the effects of the interaction between the perceived quality by the students and the outcome of formative process. Student judgements on teaching and facilities are usually collected before students give the examination but some positive and negative influences of previous experience could be taking into ac-

count to understand the results. Furthermore teacher effects on student achievement should not be undervalued.

Summing up, the evaluation of the University performance in terms of student satisfaction and achievement is a complex process that involves different level of actions and analyses, each of them to be carried out with proper methodological tools.

References

Alma Mater Studiorum Università di Bologna, Nucleo di Valutazione di Ateneo (2007), *Relazione relativa all'opinione degli studenti frequentanti in merito alle attività didattiche*.

Bartholomew D.J., Steele F., Moustaki I., Galbraith J.I. (2002), *The analysis and interpretation of multivariate data for social scientist*, Chapman and Hall, New York.

Bianconcini S., Cagnone S., Mignani S., Monari P. (2007), A latent curve analysis of unobserved heterogeneity, *Statistica*, LXVII, 40–56.

Commission of the European Communities (2007), *Progress towards the Lisbon objectives in education and training: indicators and benchmarks*, Commission Staff Working Document, http://ec.europa.eu/education/policies/2010/progressreport_en.html.

Gelmann A., Hill J. (2007), *Data analysis using regression and multilevel/hierarchical models*, Cambridge University Press, New York.

Minelli E., Rebori G., Turri M. (2005), *Valutare o misurare i risultati? il caso dell'Università*, Liuc Papers n. 165, Serie Economia aziendale 22, 2005.

MIUR CNSVU (2007), *Ottavo rapporto sullo stato del sistema universitario: rilevazione nuclei*, www.cnvsu.it.

Skrondal A., Rabe-Hesketh S. (2004), *Generalized latent variable modeling: multilevel, longitudinal and structural equation models*, Chapman and Hall/CRC, Boca Raton (FL).